Microtubule association of EML proteins and the EML4-ALK variant 3 oncoprotein require an N-terminal trimerization domain

Mark W. RICHARDS*[1], Laura O'REGAN*[1], Daniel ROTH[†], Jessica M. MONTGOMERY*, Anne STRAUBE[†], Andrew M. FRY*[2], Richard BAYLISS*[2]

*Department of Biochemistry, University of Leicester, Henry Wellcome Building, Lancaster Road, Leicester, LE1 9HN, U.K.
[†]Centre for Mechanochemical Cell Biology, Warwick Medical School, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, U.K.
[1]These authors contributed equally to this work.
[2]Correspondence may be addressed to either of these authors: email (amf5@le.ac.uk or rb308@le.ac.uk)

The structural co-ordinates reported for EML2 11-60 and EML4 6-64 I38M appear in the PDB under codes 4CGB and 4CGC respectively.

Page heading: EML trimerization domain drives MT-binding

Keywords: EML1, EML2, EML4-ALK, crystal structure, coiled-coil, microtubule

SUMMARY STATEMENT

We present crystal structures of a trimeric coiled-coil domain found in EML proteins. This trimerization domain mediates self-association and interactions between a subset of EML proteins. Microtubule-association of EML proteins requires the trimerization domain and an adjacent basic region.

ABSTRACT

Proteins of the echinoderm microtubule associated protein-like (EML) family contribute to formation of the mitotic spindle and interphase microtubule (MT) network. EML1-4 consist of WD40 repeats and an N-terminal region containing a putative coiled-coil. Recurrent gene rearrangements in non-small cell lung cancer (NSCLC) fuse *EML4* to anaplastic lymphoma kinase (*ALK*) causing expression of several oncogenic fusion variants. The fusions have constitutive ALK activity due to self-association through the EML4 coiled-coil. We have determined crystal structures of the coiled-coils from EML2 and EML4, which describe the structural basis of both EML self-association and oncogenic EML4-ALK activation. The structures reveal a trimeric oligomerization state directed by a conserved pattern of hydrophobic residues and salt bridges. We show that the trimerization domain (TD) of EML1 is necessary and sufficient for self-association. The TD is also essential for MT binding, however this property requires an adjacent basic region. These observations prompted us to investigate MT association of EML4-ALK and EML1-ABL1 fusions in which variable portions of the EML component are present. Uniquely, EML4-ALK variant 3, which includes the TD and basic region of EML4 but none of the WD40 repeats, was localized to MTs, both when expressed recombinantly and in a patient-derived NSCLC cell line (H2228). This raises the question of whether the mislocalization of ALK activity to MTs might influence downstream signalling and malignant properties of cells. Furthermore, the structure of EML4 TD may enable the development of protein-protein interaction inhibitors targeting the trimerization interface, providing a possible avenue towards therapeutic intervention in *EML4-ALK* NSCLC.

INTRODUCTION

Microtubules (MTs) are polymers of tubulin heterodimers that form a dynamic network as part of the eukaryotic cytoskeleton. They carry out many essential cellular processes including forming an intracellular transport system for RNAs, proteins and organelles and regulating cell motility and changes in cell shape [1, 2]. In mitosis, MTs form the spindle apparatus that segregates chromosomes into two equal sets. The dynamic nature of MTs - their stochastic growth and shrinkage - is crucial for their function and is regulated by MT-associated proteins (MAPs) [3].

The present work is focused on the echinoderm MT-associated protein (EMAP)-like (EML) family of MAPs. The archetypal member of this family was identified as the most abundant MAP in dividing echinoderm eggs and embryos [4]. Further members of the EML family associate with MTs and are essential for the proper formation of both the interphase MT network and the mitotic spindle [5-8]. EML proteins affect microtubule dynamics, and may act as scaffold proteins to localize mitotic kinases to MTs [5, 9, 10]. *Drosophila* and *C. elegans* each have a single EML homologue, whereas six EML proteins are present in humans of which EML1-4 have molecular weights in the 70-120 kDa range while EML5 and 6 are in excess of 200 kDa. The EML proteins contain WD40 repeats, and in EML1-4 these fall within a ~70 kDa core region that begins with a conserved ~60 amino acid sequence that has been termed the 'hydrophobic EML protein' (HELP) motif [5]. We recently determined the crystal structure of the core HELP/WD region of human EML1, which forms a tandem atypical β-propeller (TAPE) domain that binds soluble tubulin [11]. Most splice variants of EML1-4 are also predicted to contain a coiled-coil in the N-terminal region.

Rearrangements in the short arm of chromosome 2 leading to genetic fusions between *EML4* and the gene encoding anaplastic lymphoma kinase (*ALK*) occur in ~5% of cases of non-small cell lung cancer (NSCLC) [12]. Multiple variants of the *EML4-ALK* fusion have been identified in NSCLC resulting from translocations at different points within the *EML4* gene [13]. A similar genetic fusion has also been reported between *EML1* and the gene encoding another tyrosine kinase Abelson 1 (*ABL1*) in T-cell acute lymphoblastic leukaemia [14]. These fusions have potent transforming activity due to constitutive activation of the tyrosine kinase but confer addiction to the oncogene and inhibition of the tyrosine kinase induces apoptosis in transformed cells [15, 16]. Transforming activity of EML4-ALK variant 1 is dependent on the N-terminal region, and this is proposed to be due to oligomerization of ALK [12]. Indeed, this hypothesis appears to be confirmed by the identification of *EML4-ALK* variant 5, which encodes just the N-terminal region up to the end of the predicted coiled-coil [17].

In this study we determined the crystal structure of the N-terminal region of EML proteins that is required to drive the transforming capability of *EML4-ALK* fusions and reveal a trimeric state of oligomerization. We then investigated the role of this coiled-coil in MT binding in both EML proteins and oncogenic EML-kinase fusions.

MATERIALS AND METHODS

Gene manipulation

Human *EML* cDNAs were cloned by PCR from human cDNA (Clontech). Products encoding EML4 6-64 and EML2 isoform 2 11-60 were amplified by PCR and cloned into a modified version of pET30. Constructs were made comprising EML1 residues 167-815 (TAPE domain), 1-174 (N-terminus), 23-78 (TD) and 80-815 (Δ-TD). The EML1 AAAA mutant contained the mutations L59A and D61A. Cloning of *EML1-ABL1* and *EML4-ALK* variants was described previously [11]. Site-directed mutagenesis was carried out by the Quikchange procedure (Stratagene).

Protein expression and purification

His-tagged EML2 11-60 and EML4 6-64 I38M were expressed in B834 cells grown in Luria broth and SelenoMet Medium (Molecular Dimensions) respectively. The proteins were purified from clarified lysate in 50 mM HEPES pH 7.5, 500 mM NaCl using a nickel-charged 5 mL Chelating Sepharose column and eluted on a gradient from 50-500 mM imidazole. His-tags were

removed by digestion with tobacco etch virus (TEV) protease and the protein passed through a second Chelating Sepharose column. Finally, the proteins were gel filtrated into 10 mM Hepes pH 7.5, 80 mM NaCl, 2 mM dithiothreitol (DTT) on a Superdex 75 16/60 column (GE Healthcare).

Crystal structure determination

Crystals of EML4 6-64 I38M and EML2 11-60 were obtained by the hanging drop vapour diffusion method using 100 mM Tris/HCl pH 8.5, 30% PEG 4000, 200 mM $MgCl_2$ and 100 mM Bis-Tris-propane pH 6.5, 20% PEG 3350, 200 mM potassium thiocyanate respectively in the reservoir and were flash frozen in the same solutions with the addition of 20% glycerol and 5 mM DTT. Drops were formed from ~15 mg/mL protein solution mixed 1:1 with reservoir solution and incubated at 18°C for ~2 days.

Single-wavelength anomalous dispersion data to a resolution of 2.9 Å was collected from an EML4 6-64 I38M selenomethionine derivative crystal at Diamond I02. Diffraction images were indexed using imosflm [18] and scaled using scala [19]. An initial map and model were obtained using PHENIX to locate selenium sites [20] revealing one trimer in the asymmetric unit. Manual rebuilding was carried out using COOT [21]. The 2.1 Å data was collected from a native EML2 11-60 crystal at Diamond I04. PHENIX was used to solve the structure of EML2 11-60 by molecular replacement phasing with the EML4 6-64 I38M structure, to automatically build the two trimers in the asymmetric unit, and for refinement.

In vitro MT-binding assays

Tubulin was prepared from pig brain as described previously [22]. Labeled tubulin was from Cytoskeleton Inc. YFP-EML1 proteins for use in *in vitro* microtubule binding assays were expressed as FLAG-Strep2 fusions in HEK293F suspension culture and purified from clarified cell lysate on Strep-tactin Sepharose (Qiagen) in Tris-buffered saline (TBS) and eluted using 3 mM desthiobiotin (IBA), followed by incubation with GST-tagged 3C protease to cleave off the tag and then passed over a Glutathione-sepharose column to remove the protease. Finally, the protein was separated from co-purifying tubulin by gel filtration on a Superose 6 column (GE Healthcare) in 20 mM Tris/HCl pH 7.5, 80 mM NaCl, 2 mM DTT before dialysis into MRB80 (80 mM PIPES, pH 6.8 with KOH, 1 mM EGTA, 4 mM $MgCl_2$) + 50 mM KCl and concentration to 20 µM. MT seeds were assembled from tubulin, biotin-tubulin and Hilyte647-tubulin at a molar ratio of 25:1:2 in the presence of 1 mM GMP-CPP in MRB80 for 1 hr at 37ºC, diluted 20-fold with MRB80 + 2 µM Taxol and stored at 21°C. A 100 µm deep flow chamber was made from a slide and a HCl-treated coverslip using double-sided Scotch tape (3M) and passivated with PLL-PEG-50% biotin (Susos AG). Seeds were attached to this surface using streptavidin and blocked with 1 mg/mL κ-casein (Sigma). A reaction mix containing 12 µM tubulin, 1 µM rhodamine-tubulin, 50 mM KCl, 1 mM GTP, 0.6 mg/ml κ-casein, 0.2% methyl cellulose, 4 mM DTT, 0.2 mg/mL catalase, 0.4 mg/mL glucose oxidase, 50 mM glucose in MRB80, supplemented with 500 nM YFP-EML1 protein or buffer was clarified for 8 min at 190,000xg in an airfuge (Beckman), the supernatant added to the flow chamber and sealed with candle wax. For MT-binding experiments in the absence of free tubulin the reaction mix was prepared as above but lacking tubulin and assays carried out using GMP-CPP stabilized MTs.

MTs were observed on an Olympus TIRF system with a 100x NA 1.49 objective, 488 nm, 561 nm and 640 nm laser lines and a Hamamatsu ImageEM-1k back-illuminated EM-CCD camera under the control of xcellence software. To quantify MT-bound YFP-EML1 proteins the average YFP-fluorescence intensity along a 3 pixel wide line along the MT was measured using ImageJ.

Cell culture and transfection

HeLa, HEK293 and U2OS cells were grown in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% heat-inactivated foetal bovine serum (FBS), 100 IU/mL penicillin and 100 µg/mL streptomycin at 37°C in a 5% CO2 atmosphere. *EML* constructs, *EML1-ABL1* and *EML4-ALK* variants were cloned into pcDNA3 or pcDNA3.1-hygro (Invitrogen) with N-terminal YFP- or FLAG-tags. HeLa and HEK293 cells were transiently transfected with these plasmids using Fugene HD reagent (Promega) and U2OS cells using Lipofectamine 2000 (Invitrogen).

HEK293F cells were grown in suspension culture in Freestyle 293 medium (Gibco) in a 5% CO_2 atmosphere and transfected with Strep2-YFP-*EML1* constructs in a modified version of pMAX (Lonza) at a density of $1x10^6$ cells/mL using 2 µg/mL polyethylenimine.

NCI-H2228 cells were provided by Dr S Gray (Institute of Molecular Medicine, St James's Hospital, Dublin, Ireland); NCI-H3122 cells were obtained from the NCI Tumor Repository (Frederick, MD, USA). Cells were cultured in RPMI medium 1640, 2mM Glutamax and 10% Foetal Calf Serum (Gibco).

Fixed and live cell IF microscopy

For fixed cell microscopy, HeLa, H2228 and H3122 cells grown on acid-etched coverslips were fixed with ice cold methanol and processed for IF microscopy as described previously [23], with the modification that cells were blocked for 30 min in 1x PBS supplemented with 1% bovine serum albumin and 0.2% Triton X-100. Primary antibodies used were 1 µg/mL anti-GFP ab6556 (Abcam), 0.3 µg/mL anti-α-tubulin (Sigma) and 1:200 anti-ALK D5F3 (Cell Signalling). Secondary antibodies were 1 µg/mL Alexa Fluor 488 conjugated goat anti-rabbit and Alexa Fluor 594 conjugated goat anti-mouse IgGs (Invitrogen). Cells were imaged using a Leica SP5 laser scanning confocal microscope. Deconvolution of image stacks was carried out using Huygens Essential software and the z-stacks were assembled in ImageJ. Quantitative image analysis performed with LAS-AF software (Leica) or ImageJ.

Western blot analysis

Primary antibodies for western blotting were mouse anti-FLAG M2 (Sigma) and rabbit anti-GFP ab6556 (Abcam). Secondary antibodies were horseradish peroxidase conjugated anti-mouse or anti-rabbit IgGs (GE Healthcare).

RESULTS

A conserved region in the EML N-terminus forms a trimerization domain

Proteins of the EML family contain regions N-terminal to the conserved TAPE core that are poorly conserved, often subject to splice variation and predicted to be mostly disordered. However, there is an island of conserved primary sequence within this region that is consistent with a coiled-coil (Figure 1A). To investigate the structural basis of self-association we crystallized protein fragments corresponding to this part of human EML2 and EML4 proteins. The structure was determined by single-wavelength anomalous dispersion phasing using crystals of a selenomethionine-labelled EML4 6-64 I38M mutant (Table 1, Figure 1B, Supplementary Figure S1A). Only residues 14-44 were visible in the electron density map but most strikingly they formed a trimeric coiled-coil. We then obtained 2.1 Å diffraction data using crystals of EML2 isoform 2 residues 11-60 and solved its structure by molecular replacement phasing (Table 1, Figure 1C, Supplementary Figure S1B).

The EML trimerization domain (TD) forms an amphipathic α-helix with a series of nine conserved hydrophobic residues (predominantly leucine and valine) along the inner face whose side chains pack together into the core of the trimer (coloured yellow in Figure 1). A conserved glutamine (Gln29 in EML2, coloured grey in Figure 1D) on the inner face makes hydrogen bond interactions with main-chain carbonyls in the core, perhaps serving to maintain the register of the coiled-coil. Interspersed between the hydrophobic residues in the sequence are conserved charged residues whose side chains are positioned to form salt-bridges with those of the adjacent protomers around the outside of the trimer (coloured light blue in Figure 1). The pattern of conserved residues indicates that human EML1-4 are all trimeric, as are their respective oncogenic fusion proteins (Figure 1E).

The trimerization domain is critical for EML self-association

The most conserved surface patch in the TD is composed of several amino acids with short side chains (ALAD motif), which constrict the shape of the coiled-coil (Figure 2A). We investigated whether the TD or disruption of the ALAD motif affect self-association in a representative member of the EML family, EML1. We tested the ability of various YFP-tagged EML1 constructs to co-immunoprecipitate with FLAG-tagged full-length EML1 from lysates of co-transfected HEK293 cells (Figure 2B). Full-length EML1 (FL), the isolated N-terminus (1-174) and a construct corresponding to residues 23-78 (trimerization domain; TD) co-immunoprecipitated with full-length EML1 while the TAPE region did not. A double mutation (L59A, D61A) in the most conserved motif in the TD (AAAA) did not disrupt self-association but deletion of the whole conserved region (Δ-TD) abrogated self-association completely.

The key residues involved in self-association within the trimerization domain are imperfectly conserved, and EML3 has the most divergent TD (Figure 1E). This observation prompted us to investigate whether EML3 could self-associate, and whether associations between EML3 and other members of the EML family could happen. We tested the ability of YFP-EML3 to co-immunoprecipitate with FLAG-EML1, FLAG-EML2 or FLAG-EML3 from lysates of co-transfected HEK293 cells (Figure 2C). YFP-EML3 co-immunoprecipitated with FLAG-EML3 and, more weakly, with FLAG-EML2, but not with YFP-EML1. This is consistent with the degree of conservation within the 17 key residues involved in self-association: EML1 and EML2 - 12 identical, 5 conservatively substituted; EML1 and EML3 - 6 identical, 5 conservatively substituted, 5 non-conserved; EML2 and EML3 - 9 identical, 3 conservatively substituted, 5 non-conserved.

The trimerization domain is critical for MT binding

We confirmed co-localization of YFP-tagged human EML1 to the interphase MT network of transiently transfected HeLa cells, as has been previously observed for endogenous EML proteins (Figure 3) [5-8]. Overexpression of a tagged protein might affect its regulation and distribution on microtubules, and we shall not comment on these features of the data. However, this approach enabled us to screen the MT-binding properties of multiple YFP-EML1 constructs to identify which domains were needed for this activity (Figure 3A). The co-localization of YFP-EML1 proteins with MTs was visualised (Figure 3B) and quantified (Figure 3C).

The isolated N-terminal region (aa 1-174), which is not conserved at the sequence level for most of its length but includes the TD, strongly associated with MTs. However, neither the putatively disordered region lying between the TD and TAPE domains (74-174) nor the TD alone (23-78), nor the isolated TAPE domain were sufficient to associate with MTs in cells. Deletion of the TD from EML1 considerably reduced its interaction with MTs in cells, whereas mutation of the ALAD motif within the TD decreased MT co-localization modestly. The overexpressed EML1 proteins might interact directly with MTs, through association with endogenous EML1 or other EMLs, or indeed through associations with other MAPs. To distinguish between these different modes of interaction, direct binding of recombinantly expressed and purified YFP-EML1 protein constructs to purified MTs was measured *in vitro* by total internal reflection fluorescence (TIRF) microscopy (Figure 4). Here, direct association of full-length EML1 and the N-terminal region with MTs was measured, whereas the isolated TAPE domain did not associate (Figure 4A,C); full-length EML1 with a mutated ALAD motif exhibited reduced binding compared to wild-type, and the deletion of the TD abolished MT interaction altogether (Figure 4B,D).

Taken together, the results from the cell-based and *in vitro* MT-binding assays demonstrate that the N-terminal region of EML1 confers MT binding and that the TD is required. However, the TD is not sufficient for MT association, which requires other sequences in the N-terminal region.

EML4-ALK variant 3 localizes to MTs

Having identified the N-terminus as a MT-binding region of EML proteins, and recognising that it is intact in many of the fusion oncoproteins, we supposed that the fusion oncoproteins might be targeted to MTs. We examined HeLa cells transfected with YFP-tagged EML-tyrosine kinase

fusions (Figure 5A-C) and found that this was not generally the case: EML4-ALK variants 1 and 2 and EML1-ABL1 form cytoplasmic aggregates and EML4-ALK variant 5a is diffusely localized in the cytoplasm, while only EML4-ALK variant 3a was localized to MTs. Next we investigated by immunofluorescence microscopy the localization of the endogenous EML4-ALK fusion proteins in two patient-derived NSCLC cell lines, H3122 and H2228 (Figure 5D-F). The variant 1 protein present in H3122 cells displayed a patchy appearance in the cytoplasm consistent with a protein that is largely aggregated but the variant 3b protein in H2228 cells was localized to the MT network.

Taken together, these results show that the N-terminal domain of EML4 is able to confer MT localization on EML4-ALK variant 3, but not when a truncated part of the EML4 TAPE domain is present as in variants 1 and 2.

DISCUSSION

Association of the EML1 N-terminal region with MTs requires both the TD and the ~90 residue region between the TD and TAPE domains. This latter, putatively disordered, inter-domain region is not conserved at the sequence level but contains many basic residues. The TD presumably enhances binding of EML proteins to MTs through an avidity effect, and may contribute directly to the interaction through a conserved surface region that is disrupted by the AAAA mutation.

EML oncogenic fusions such as EML4-ALK and EML1-ABL1 are naturally occurring and exist in a context in which the MT binding properties of the EML N-terminus are isolated from the TAPE region. In EML4-ALK variant 3a/b the whole of the EML4 N-terminal region, and evidently its MT-binding activity, is transferred to the fusion protein (Figure 6). EML4-ALK variant 5 has a TD, but lacks the basic region, and is therefore incapable of binding to MTs. The other EML4-ALK variants and EML1-ABL1 contain a complete N-terminus encompassing the TD and basic region, and might be expected to localize to MTs, but do not. In these cases, it appears that the presence of a partial TAPE domain interferes with MT binding, perhaps through protein misfolding leading to aggregation into complexes with Hsp90 that may sterically block MT association. Indeed, the microscopy studies of several of the EML-kinase fusions, including endogenous EML4-ALK variant 1 in H3122 cells, were consistent with aggregation of these proteins. Furthermore, we have previously shown that EML4-ALK variants 1 and 2 are more sensitive to Hsp90 inhibitors than variants 3 and 5 [11].

The EML2 and EML4 TD structures presented here reveal that EML proteins and EML4-ALK oncoproteins self-associate as trimers through a specific and stable coiled-coil interface. The more widely expressed isoform 1 of EML2 does not possess a TD region and is therefore unlikely to interact directly with other EMLs but a TD is present in the brain/spinal cord variant isoform 2 studied here [24]. EML3 can also self-associate, and can interact with EML2 isoform 2, but not with EML1. The functional relevance of interactions between EML family members is not clear, but the potential for these interactions should be borne in mind during functional studies.

It is notable that the more divergent human paralogues, EML5 and EML6, lack coiled-coil regions but have three copies of the TAPE region each. Thus, the quaternary structure of all six human EML proteins may feature three TAPE regions (Figure 6A). In oncogenic ALK-fusions oligomerization conferred by the fusion partner mimics the native oligomerization-dependent activation mechanism of receptor tyrosine kinases and engenders constitutive activation of ALK [25]. The use of rationally-designed peptide and peptidomimetic inhibitors is under investigation for the disruption of several coiled-coil interactions that are implicated in disease and clinically useful therapeutics have been developed using this approach [27, 28, 29]. The TD structures may enable the development of protein-protein interaction inhibitors targeting the trimerization interface of the EML4 TD, another possible avenue towards therapeutic intervention in *EML4-ALK* NSCLC and one that would be applicable to all variants.

AUTHOR CONTRIBUTION

Mark Richards made the DNA constructs, protein samples and crystals and determined the crystal structures. Laura O'Regan carried out immunofluorescence microscopy and analysed the resulting data, while immunoprecipitation experiments were performed and analysed by Laura O'Regan and Jessica Montgomery. Daniel Roth and Anne Straube carried out *in vitro* MT-binding experiments and analysed the resulting data. Mark Richards, Richard Bayliss, Anne Straube and Andrew Fry wrote the manuscript.

FUNDING

This work was supported by a Royal Society University Research Fellowship and Cancer Research UK programme funding to R.B. [C24461/A12772], a Marie Curie Cancer Care programme grant and Lister Institute Research Prize to A.S. and funding from Worldwide Cancer Research and the Wellcome Trust to A.M.F.

ACKNOWLEDGEMENTS

We thank the beamline support staff at DLS I02 and I04 for assistance with data collection and the Centre for Core Biotechnology Services (University of Leicester) for imaging support.

REFERENCES

1 Hirokawa, N., Noda, Y., Tanaka, Y. and Niwa, S. (2009) Kinesin superfamily motor proteins and intracellular transport. Nat. Rev. Mol. Cell Biol. **10**, 682-696

2 Kaverina, I. and Straube, A. (2011) Regulation of cell migration by dynamic microtubules. Semin. Cell Dev. Biol. **22**, 968-974

3 van der Vaart, B., Akhmanova, A. and Straube, A. (2009) Regulation of microtubule dynamic instability. Biochem. Soc. Trans. **37**, 1007-1013

4 Suprenant, K. A., Dean, K., McKee, J. and Hake, S. (1993) EMAP, an echinoderm microtubule-associated protein found in microtubule-ribosome complexes. J. Cell Sci. **104**, 445-450

5 Eichenmuller, B., Everley, P., Palange, J., Lepley, D. and Suprenant, K. A. (2002) The human EMAP-like protein-70 (ELP70) is a microtubule destabilizer that localizes to the mitotic apparatus. J. Biol. Chem. **277**, 1301-1309

6 Houtman, S. H., Rutteman, M., De Zeeuw, C. I. and French, P. J. (2007) Echinoderm microtubule-associated protein like protein 4, a member of the echinoderm microtubule-associated protein family, stabilizes microtubules. Neuroscience. **144**, 1373-1382

7 Pollmann, M., Parwaresch, R., Adam-Klages, S., Kruse, M. L., Buck, F. and Heidebrecht, H. J. (2006) Human EML4, a novel member of the EMAP family, is essential for microtubule formation. Exp. Cell. Res. **312**, 3241-3251

8 Tegha-Dunghu, J., Neumann, B., Reber, S., Krause, R., Erfle, H., Walter, T., Held, M., Rogers, P., Hupfeld, K., Ruppert, T., Ellenberg, J. and Gruss, O. J. (2008) EML3 is a nuclear microtubule-binding protein required for the correct alignment of chromosomes in metaphase. J. Cell Sci. **121**, 1718-1726

9 Brisch, E., Daggett, M. A. and Suprenant, K. A. (1996) Cell cycle-dependent phosphorylation of the 77 kDa echinoderm microtubule-associated protein (EMAP) in vivo and association with the p34cdc2 kinase. J Cell Sci. **109 (Pt 12)**, 2885-2893

10 Hamill, D. R., Howell, B., Cassimeris, L. and Suprenant, K. A. (1998) Purification of a WD repeat protein, EMAP, that promotes microtubule dynamics through an inhibition of rescue. J. Biol. Chem. **273**, 9285-9291

11 Richards, M. W., Law, E. W., Rennalls, L. P., Busacca, S., O'Regan, L., Fry, A. M., Fennell, D. A. and Bayliss, R. (2014) Crystal structure of EML1 reveals the basis for Hsp90 dependence of oncogenic EML4-ALK by disruption of an atypical beta-propeller domain. Proc. Natl. Acad. Sci. U.S.A. **111**, 5195-5200

12 Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y. and Mano, H. (2007) Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. **448**, 561-566

13 Sasaki, T., Rodig, S. J., Chirieac, L. R. and Janne, P. A. (2010) The biology and treatment of EML4-ALK non-small cell lung cancer. Eur. J. Cancer. **46**, 1773-1780

14 De Keersmaecker, K., Graux, C., Odero, M. D., Mentens, N., Somers, R., Maertens, J., Wlodarska, I., Vandenberghe, P., Hagemeijer, A., Marynen, P. and Cools, J. (2005) Fusion of EML1 to ABL1 in T-cell acute lymphoblastic leukemia with cryptic t(9;14)(q34;q32). Blood. **105**, 4849-4852

15 Choi, Y. L., Takeuchi, K., Soda, M., Inamura, K., Togashi, Y., Hatano, S., Enomoto, M., Hamada, T., Haruta, H., Watanabe, H., Kurashina, K., Hatanaka, H., Ueno, T., Takada, S.,

Yamashita, Y., Sugiyama, Y., Ishikawa, Y. and Mano, H. (2008) Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer. Cancer Res. **68**, 4971-4976
16 Takezawa, K., Okamoto, I., Nishio, K., Janne, P. A. and Nakagawa, K. (2011) Role of ERK-BIM and STAT3-survivin signaling pathways in ALK inhibitor-induced apoptosis in EML4-ALK-positive lung cancer. Clin. Cancer Res. **17**, 2140-2148
17 Takeuchi, K., Choi, Y. L., Soda, M., Inamura, K., Togashi, Y., Hatano, S., Enomoto, M., Takada, S., Yamashita, Y., Satoh, Y., Okumura, S., Nakagawa, K., Ishikawa, Y. and Mano, H. (2008) Multiplex reverse transcription-PCR screening for EML4-ALK fusion transcripts. Clin. Cancer Res. **14**, 6618-6624
18 Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D Biol. Crystallogr. **67**, 271-281
19 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50**, 760-763
20 Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. **58**, 1948-1954
21 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**, 2126-2132
22 Gell, C., Friel, C. T., Borgonovo, B., Drechsel, D. N., Hyman, A. A. and Howard, J. (2011) Purification of tubulin from porcine brain. Methods Mol. Biol. **777**, 15-28
23 O'Regan, L. and Fry, A. M. (2009) The Nek6 and Nek7 protein kinases are required for robust mitotic spindle formation and cytokinesis. Mol. Cell. Biol. **29**, 3975-3990
24 Lepley, D. M., Palange, J. M. and Suprenant, K. A. (1999) Sequence and expression patterns of a human EMAP-related protein-2 (HuEMAP-2). Gene. **237**, 343-349
25 Roskoski, R., Jr. (2013) Anaplastic lymphoma kinase (ALK): structure, oncogenic activation, and pharmacological inhibition. Pharmacol. Res. **68**, 68-94
26 Hames, R. S., Crookes, R. E., Straatman, K. R., Merdes, A., Hayes, M. J., Faragher, A. J. and Fry, A. M. (2005) Dynamic recruitment of Nek2 kinase to the centrosome involves microtubules, PCM-1, and localized proteasomal degradation. Mol. Biol. Cell. **16**, 1711-1724
27 Wild, C.T., Shugars, D.C., Greenwell, T.K., McDanal, C.B. and Matthews, T.J. (1994) Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. Proc. Natl. Acad. U.S.A. **91**, 9770-9774
28 Dixon, A.S., Pendley, S.S., Bruno, B.J., Woessner, D.W., Shimpi, A.A., Cheatham III, T.E and Lim, C.S. (2011) Disruption of Bcr-Abl coiled-coil oligomerization by design. J. Biol. Chem. **286**, 27751-27760
29 Rao, T., Ruiz-Gómez, G., Hill, T.A., Hoang, H.N., Fairlie, D.P., Mason, J.M. (2013) Truncated and helix-constrained peptides with high affinity and specificity for the coiled-coil of AP-1. PLoS ONE **8**, e59415-59415

Figure Legends

Figure 1 The N-termini of EML proteins contain a trimerization domain (TD). (**A**) Schematic diagram of domain organization in EML1. (**B**) A view along the threefold axis of the EML4 TD. In B-D, side chains forming the hydrophobic core of the trimer are shown in yellow and those forming salt-bridges are shown in light blue. (**C**) A view along the threefold axis of the EML2 TD. (**D**) The structure of the EML2 TD shown as a cartoon representation (pink) with key residues in two of the protomers shown as stick representations. Gln29 is shown in grey. Below, the arrows linking the pair of sequences indicate salt-bridge pairing in adjacent protomers. (**E**) A sequence alignment of the TD region of human EML proteins. Residues are coloured as in B-D and boxes indicate the regions present in the EML2 and EML4 TD structures. Conservation is shown underneath - identical residues and highly conserved residues are marked by (*) and (:) respectively. Gln29 and the ALAD motif are underlined.

Figure 2 Oligomerization of EML1 depends on the TD. (**A**) Sequence conservation mapped onto the surface of the EML2 TD structure; residues that are identical among human EML proteins are coloured red and highly conserved residues are coloured orange. (**B**) Immunoblot analysis of EML1 self-association. HEK293 cells were co-transfected with FLAG-*EML1* and various YFP-*EML1* constructs as indicated (top two panels). Proteins were immunoprecipitated using anti-FLAG M2 (Sigma) (lower two panels). Cell extracts were prepared as previously described [26] and analysed by SDS-PAGE and western blotting using anti-FLAG M2 and anti-GFP antibodies. (**C**) EML3 association with EML1-3. U2OS cells were co-transfected with YFP-*EML3 and* FLAG-*EML1-3* as indicated. Immunoprecipitation and analysis was carried out as in (B).

Figure 3 Microtubule association of EML1 (**A**) Schematic illustrations of EML1 constructs. (**B**) Immunofluorescence microscopy images of interphase HeLa cells transfected with YFP-tagged *EML1* constructs (z-projections; lower image) and of a magnified single MT (one focal plane; upper image). Cells are stained with Hoechst 33258 (blue, DNA) and anti-GFP (green) and anti-α-tubulin (red, MTs) antibodies. Scale bar, 10 μm. (**C**) Representative intensity profiles along a line crossing 2-3 MTs showing co-localization of YFP-fusions with MTs. R (Pearson's correlation coefficient) measures the correlation between the YFP and α-tubulin signals and is given as the mean of single line-scan measurements in 10 different cells for each construct +/- standard deviation (SD).

Figure 4 Association of EML1 with microtubules assembled from purified tubulin (A) Representative TIRF microscopy fields of view showing Hilyte640-labelled GMPCPP-stabilised MT seeds (left), XRhodamine-labelled dynamic MTs extensions (centre) and the binding to those MTs of 500 nM YFP-EML1 proteins as indicated in the presence of 12 μM free tubulin (right). Note even decoration with no preference for nucleotide state of tubulin. (B) TIRF images of Hilyte640-labelled GMPCPP-stabilised MTs (left) and their decoration by 500 nM YFP-tagged EML1 mutants in the absence of free tubulin (right). Scale bars 10μm. (**C**) Statistical box plots describing quantification of MT-binding of YFP-EML1 proteins *in vitro* by TIRF microscopy in the presence of 500 nM YFP-tagged EML1 domains and 12 μM free tubulin. n=25-38 MTs. (**D**) Statistical box plots describing quantification of MT-binding of 500 nM YFP-tagged EML1 mutants without free tubulin. n= 30-201 microtubules. Boxes in C and D show 25th, 50th and 75th percentile, whiskers show 10th and 90th percentile of data.

Figure 5 Localization of EML4-ALK and EML1-ABL1 fusion proteins. (**A**) Schematic illustrations of EML-tyrosine kinase fusion constructs. (**B**) Immunofluorescence microscopy images of interphase HeLa cells transfected with YFP-tagged EML-tyrosine kinase fusion constructs as indicated, (lower image) and of a magnified single MT (upper image). Scale bar, 10 μm. (**C**) Co-localization of each YFP-fusion with MTs was quantified as for Figure 3C. (**D**) Schematic illustrations of EML-tyrosine kinase fusion constructs in H3122 and H2228. (**E**) Immunofluorescence microscopy images of interphase NSCLC cells (lower image) and of a

magnified single MT (upper image). Scale bar, 10 μm. (**F**) Co-localization of ALK with MTs was quantified as for Figure 3C.

Figure 6 Models summarising the properties of EML proteins and EML4-ALK variants. (**A**) Cartoon illustrating trimerization and MT-binding in EML proteins. (**B**) Cartoon illustrating trimerization and MT binding in EML4-ALK variant 3. The constituent β-propellers of the EML TAPE domain are shown in teal and green and the Blade 12-N subdomain is shown in orange [11].

Supplementary Figure S1 Stereo images showing representative electron density maps for EML trimerization domain X-ray crystal structures. Wire mesh shows 2mFo-DFc map contoured at 1.0σ. (**A**) The 2.9Å structure of the EML4 trimerization domain I38M showing inter-protomer salt-bridges. (**B**) The 2.1Å structure of the EML2 trimerization domain showing the hydrophobic core.

Table 1 Summary of crystallographic analysis

	EML4 TD I38M	EML2 TD
Lattice parameters		
Space group	$P4_3\,2_1\,2$	$P2_1\,2_1\,2_1$
a (Å)	67.49	47.87
b (Å)	67.49	51.14
c (Å)	50.60	101.69
α (°)	90.00	90.00
β (°)	90.00	90.00
γ (°)	90.00	90.00
Data collection		
X-ray source	Diamond I02	Diamond I04
Wavelength (Å)	0.9796	0.9795
Resolution range (Å)	67.49-2.90 (3.06-2.90)*	50.86-2.15 (2.32-2.15)
No. unique reflections	2864 (394)	14593 (1058)
Completeness (%)	100.0 (100.0)	99.8 (99.9)
Redundancy	13.4 (14.4)	7.0 (7.3)
R_{merge} (%)	13 (125)	8.5 (63.8)
R_{pim} (%)	4.1 (33.2)	3.8 (27.2)
I / σI	13.3 (3.1)	14.2 (3.2)
Phasing (figures of merit)		
before density modification	0.39	-
after density modification	0.74	-
Refinement		
Resolution range (Å)	47.7-2.9	50.9-2.2
R_{work} / R_{free} (%)	26.87 / 29.15	20.78 / 25.15
No. protein molecules	3	6
No. amino acids	83	254
No. water molecules	0	102
No. hetero molecules	0	3
Mean B-factors (Å^2)		
protein	88.75	43.30
water	-	46.05
hetero	-	50.31
r.m.s. deviations		
bond lengths (Å)	0.006	0.006
bond angles (°)	1.02	0.930
MolProbity analysis		
All-atom clash-score	17.12	8.97
Poor rotamers (%)	8.8	5.1
Ramachandran outliers (%)	0.0	0.0
Ramachandran favoured (%)	100.0	99.6
MolProbity score	1.99	1.93

*Values in parentheses are for the highest-resolution shell.

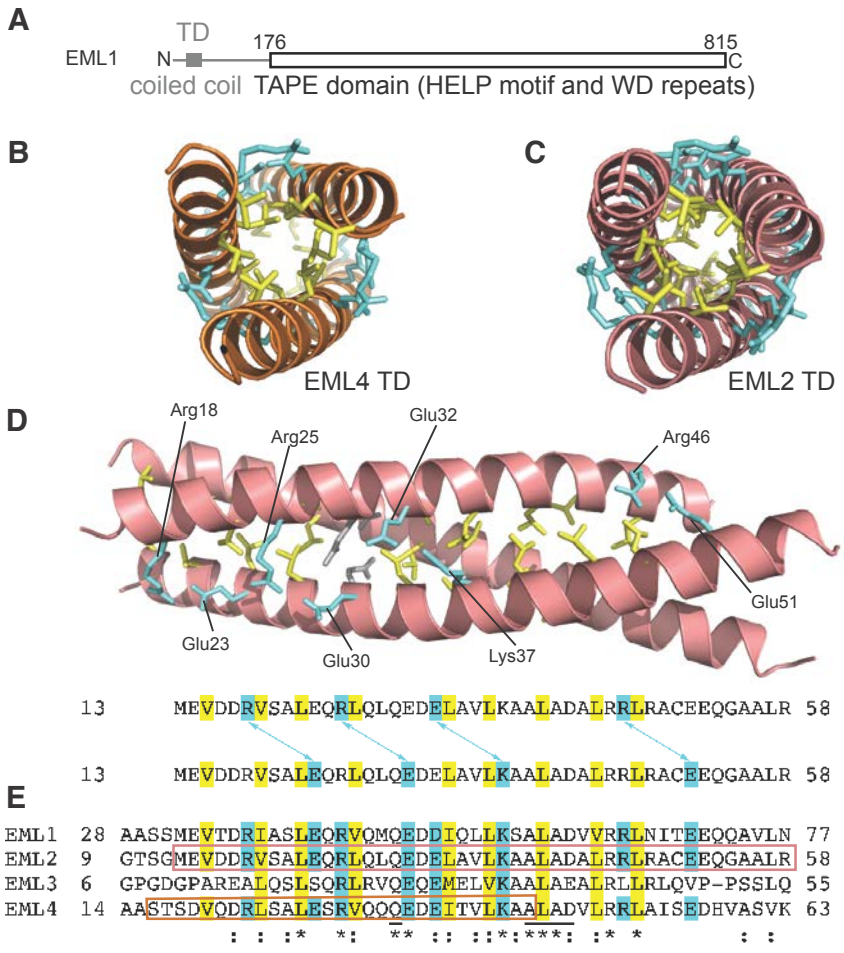

A
TD
EML1
N
176
815
C
coiled coil
TAPE domain (HELP motif and WD repeats)
B
EML4 TD
C
EML2 TD
D
Arg18
Arg25
Glu32
Arg46
Glu51
Glu23
Glu30
Lys37
13 MEVDDRVSALEQRLQLQEDELAVLKAALADALRRLRACEEQGAALR 58
13 MEVDDRVSALEQRLQLQEDELAVLKAALADALRRLRACEEQGAALR 58
E
EML1 28 AASSMEVTDRIASLEQRVQMQEDDIQLLKSALADVVRRLNITEEQQAVLN 77
EML2 9 GTSGMEVDDRVSALEQRLQLQEDELAVLKAALADALRRLRACEEQGAALR 58
EML3 6 GPGDGPAREALQSLSQRLRVQEQEMELVKAALAEALRLLRLQVP-PSSLQ 55
EML4 14 AASTSDVQDRLSALESRVQQQEDEITVLKAALADVLRRLAISEDHVASVK 63

Figure 1

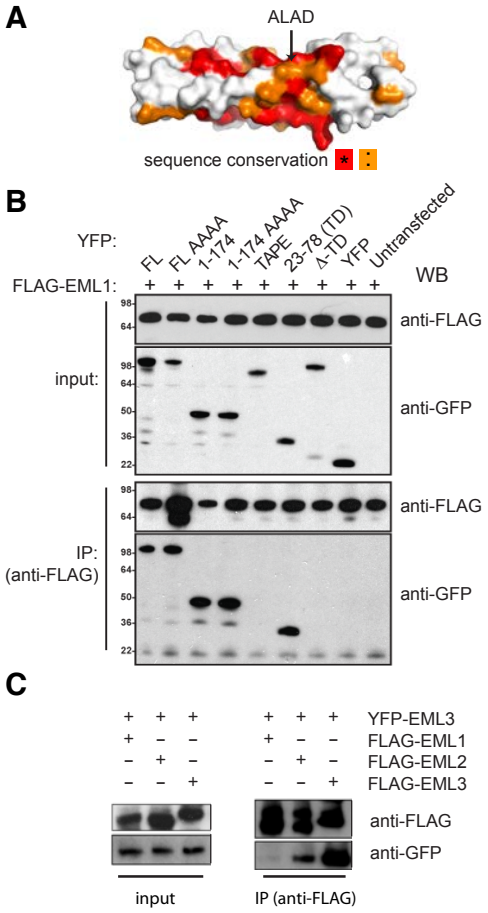

A
ALAD
sequence conservation
B
YFP:
FL
FL AAAA
1-174
1-174 AAAA
TAPE
23-78 (TD)
Δ-TD
YFP
Untransfected
WB
FLAG-EML1:
anti-FLAG
input:
anti-GFP
IP:
(anti-FLAG)
anti-FLAG
anti-GFP
C
YFP-EML3
FLAG-EML1
FLAG-EML2
FLAG-EML3
anti-FLAG
anti-GFP
input
IP (anti-FLAG)

Figure 2

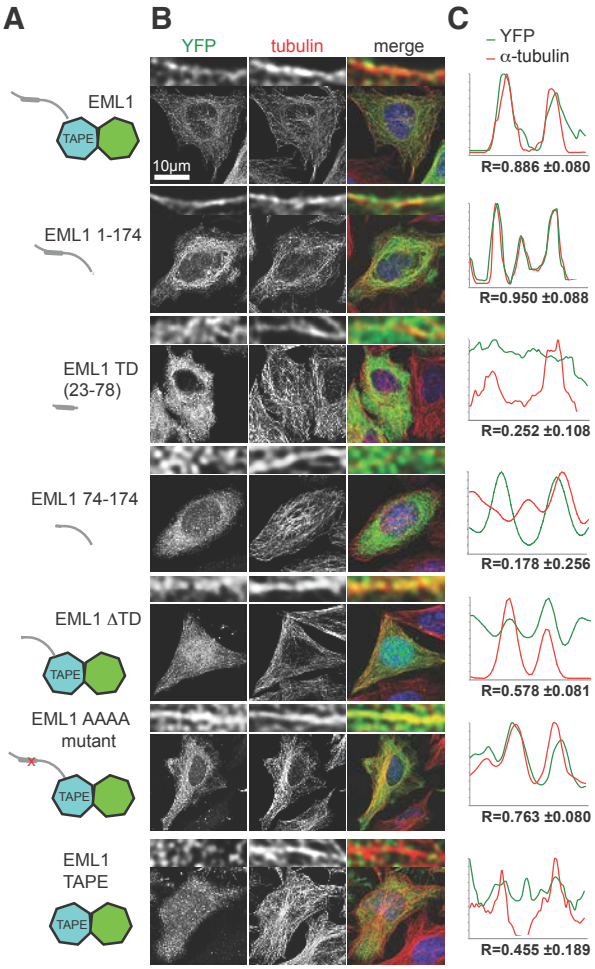

A
B
C
YFP
tubulin
merge
YFP
α-tubulin
EML1
TAPE
10μm
R=0.886 ±0.080
EML1 1-174
R=0.950 ±0.088
EML1 TD
(23-78)
R=0.252 ±0.108
EML1 74-174
R=0.178 ±0.256
EML1 ΔTD
TAPE
R=0.578 ±0.081
EML1 AAAA
mutant
TAPE
R=0.763 ±0.080
EML1
TAPE
TAPE
R=0.455 ±0.189

Figure 3

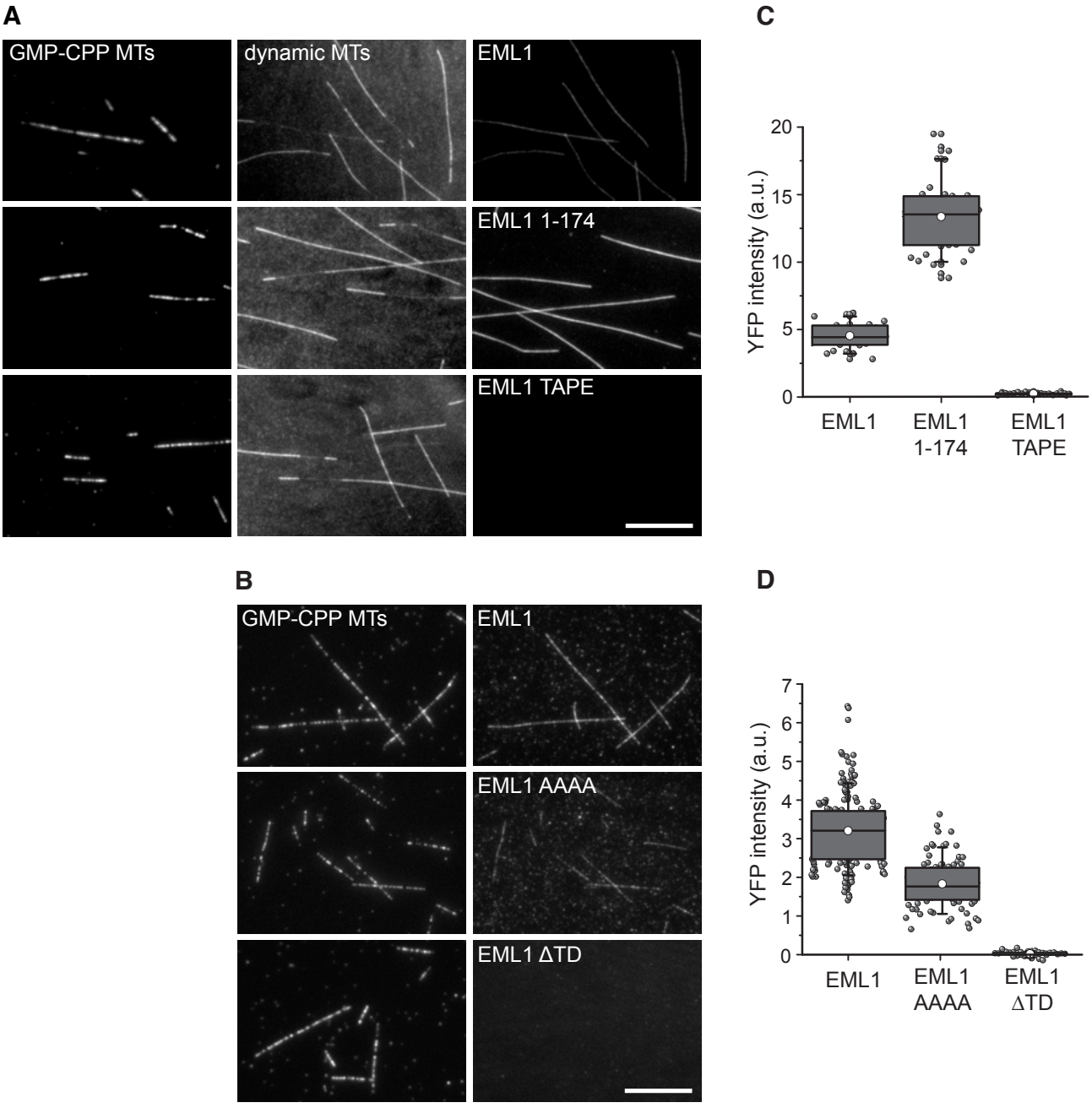

A
GMP-CPP MTs
dynamic MTs
EML1
EML1 1-174
EML1 TAPE
B
GMP-CPP MTs
EML1
EML1 AAAA
EML1 ΔTD
C
YFP intensity (a.u.)
0
5
10
15
20
EML1
EML1 1-174
EML1 TAPE
D
YFP intensity (a.u.)
0
1
2
3
4
5
6
7
EML1
EML1 AAAA
EML1 ΔTD

Figure 4

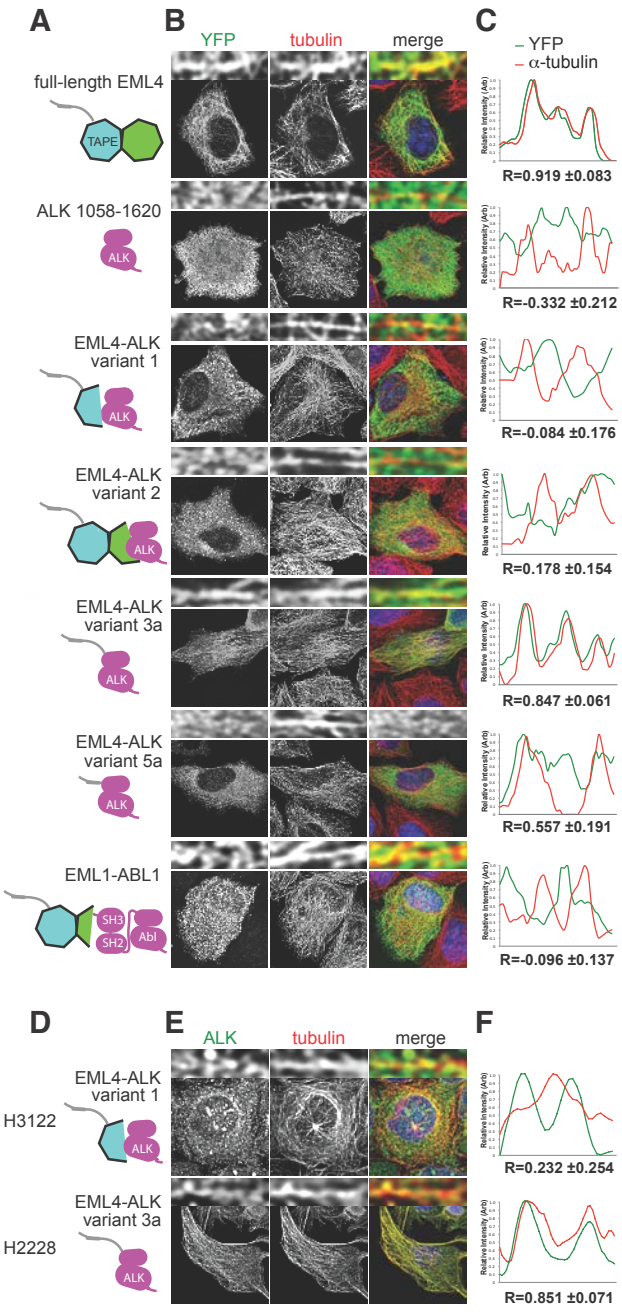

A
B
C
YFP
tubulin
merge
YFP
α-tubulin
full-length EML4
TAPE
R=0.919 ±0.083
ALK 1058-1620
ALK
R=-0.332 ±0.212
EML4-ALK variant 1
ALK
R=-0.084 ±0.176
EML4-ALK variant 2
ALK
R=0.178 ±0.154
EML4-ALK variant 3a
ALK
R=0.847 ±0.061
EML4-ALK variant 5a
ALK
R=0.557 ±0.191
EML1-ABL1
SH3
SH2
Abl
R=-0.096 ±0.137
D
E
F
ALK
tubulin
merge
EML4-ALK variant 1
H3122
ALK
R=0.232 ±0.254
EML4-ALK variant 3a
H2228
ALK
R=0.851 ±0.071

Figure 5

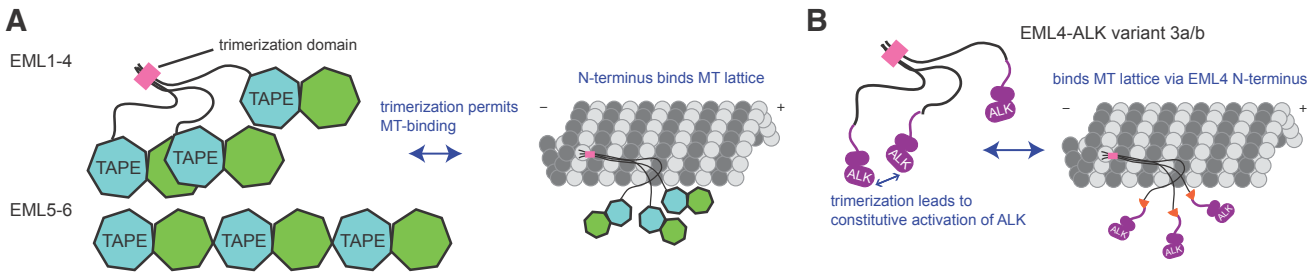
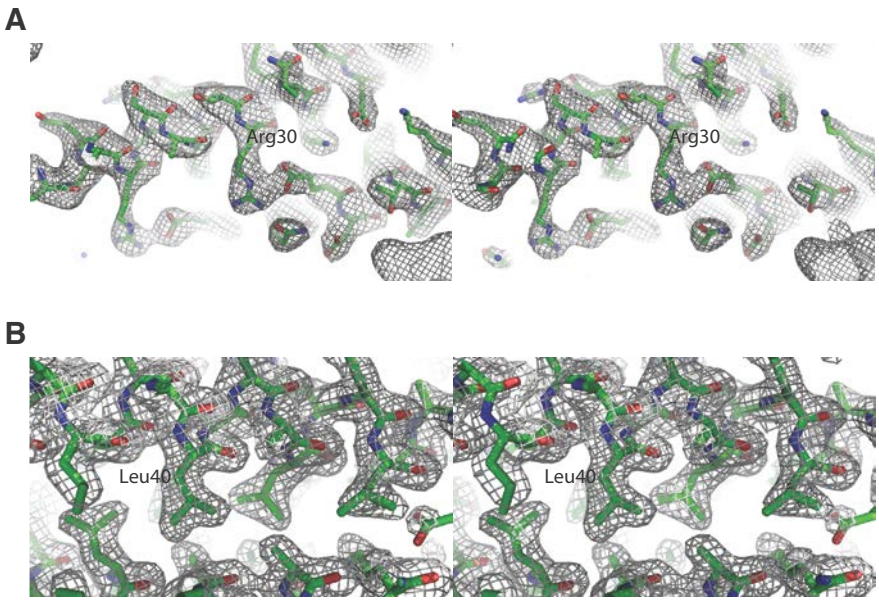

A
EML1-4
trimerization domain
TAPE
TAPE
TAPE
EML5-6
TAPE
TAPE
TAPE
trimerization permits
MT-binding
N-terminus binds MT lattice
B
EML4-ALK variant 3a/b
ALK
ALK
ALK
trimerization leads to
constitutive activation of ALK
binds MT lattice via EML4 N-terminus

Figure 6



Supplementary Figure S1